

WHAT IS CLAIMED IS:

1. A method for wholly or partially deducing the sequence of a target subunit sequence, comprising:

providing the sequences of a plurality of variants, which are distinguished individually from the target subunit sequence by the addition, deletion and/or substitution of at least one subunit;

comparing the individual sequences of said variants with each other and optionally with a sequence derived from the target subunit sequence or from a sequence adjacent thereto to deduce a consensus sequence, which corresponds to all or part of the target subunit sequence.

2. The method of claim 1, wherein the comparison is effected using sequence alignment or probabilistic techniques.

3. The method of claim 1, wherein said target subunit sequence is refractory to sequence analysis.

4. The method of claim 1, wherein said subunit sequences are selected from amino acid sequences or nucleic acid sequences.

5. The method of claim 1, wherein said subunit sequences are selected from nucleic acid sequences.

6. A method for wholly or partially determining the sequence of a primary subunit sequence, said method comprising:

providing at least one secondary subunit sequence which varies from said primary sequence by the addition, deletion and/or substitution of at least one subunit; and

comparing overlapping subsequences corresponding to said at least one secondary subunit sequence and to said primary subunit sequence to reconstruct at least a portion of the primary subunit sequence.

7. The method of claim 6, wherein the overlapping subsequences have a length p and the primary subunit sequence comprises repeated sequences of $p-1$.

8. The method of claim 7, wherein the variation in the at least one secondary subunit sequence is associated with the alteration or destruction of at least one of said repeated subsequences.

9. The method of claim 6, wherein said comparison comprises alternately reconstructing said primary subunit sequence and said at least one secondary subunit sequence using an end portion of a respective reconstruction as a guide to extend another reconstruction.

10. The method of claim 9, wherein said alternate reconstruction comprises:
comparing a portion of the primary subunit sequence with subsequences corresponding to said at least one secondary subunit sequence to identify a subsequence which aligns best with said portion and which extends unambiguously in said alignment a reconstruction of said at least one secondary subunit sequence beyond said portion; and

comparing an end portion of said reconstruction with subsequences corresponding to said primary subunit sequence to identify a subsequence which aligns best with said end portion of said reconstruction and which extends unambiguously in said alignment the reconstruction of said primary subunit sequence.

11. The method of claim 9, wherein said alternate reconstruction comprises deducing a best alignment between a subsequence and a sequence reconstruction by comparing the alignment of different subsequences with said reconstruction to produce a plurality of extended reconstructions together with individual alignment scores for each reconstruction, and optionally iteratively comparing downstream alignments of extended reconstructions using subsequences available for reconstruction, and determining a reconstruction with the highest scoring alignment to thereby deduce said best alignment.

12. The method of claim 6, further comprising generating subsequences using a sequence analysis technique.

13. The method of claim 6, wherein the sequence analysis technique is selected from a Sequence by Hybridization (SBH) technique or a shotgun sequencing technique.

14. The method of claim 6, wherein the at least one secondary subunit sequence is produced by mutagenesis of the primary subunit sequence.

15. The method of claim 6, wherein a secondary subunit sequence is produced by mutagenesis of another secondary subunit sequence.

16. The method of claim 14, wherein a parent subunit sequence is mutagenized to produce at least one variant subunit sequence in which at least 5% of subunits are different relative to the parent subunit sequence.

17. The method of claim 15, wherein a parent subunit sequence is mutagenized to produce at least one variant subunit sequence in which at least 5% of subunits are different relative to the parent subunit sequence.

18. The method of claim 6, wherein said subunit sequences are selected from amino acid sequences or nucleic acid sequences.

19. The method of claim 6, wherein said subunit sequences are selected from nucleic acid sequences.

20. The method of claim 19, wherein a parent nucleic acid sequence is mutagenized by incorporation of nucleotide analogues.

21. The method of claim 19, wherein a parent nucleic acid sequence is mutagenized by incorporation of nucleotide analogues, which are selected from dPTP (6-(2-deoxy-B-D-ribofuranosyl)-3,4-dihydro-8H-pyrimido-[4,5-C]oxazin-7-one triphosphate) or 8-oxo-dGTP (8-oxo-deoxyguanosine triphosphate).

22. The method of claim 19, wherein a parent nucleic acid sequence is mutagenized using low fidelity nucleic acid amplification reaction and an error prone DNA polymerase.

23. The method of claim 19, wherein a parent nucleic acid sequence is mutagenized using a repair deficient host.

24. A computer program product for wholly or partially deducing the sequence of a target subunit sequence, the computer program product including computer executable code which when implemented on a suitable processing system causes the processing system to perform the method of claim 1.

25. A computer program product for wholly or partially deducing the sequence of a target subunit sequence, the computer program product including computer executable code

which when implemented on a suitable processing system causes the processing system to perform the method of claim 6.

26. A computer program product stored on a computer readable medium, the computer program product being adapted to cause a processing system to wholly or partially deduce the sequence of a target subunit sequence, the computer program product including computer executable code which when implemented on a suitable processing system causes the processing system to perform the method of claim 1.

27. A computer program product stored on a computer readable medium, the computer program product being adapted to cause a processing system to wholly or partially deduce the sequence of a target subunit sequence, the computer program product including computer executable code which when implemented on a suitable processing system causes the processing system to perform the method of claim 6.